The Impact of Web Services at the IRIS DMC

Tim Ahern, Robert Weekly, Chad Trabant, Mick Van Fossen, Yanzan Suleiman, Bruce Weertman, Mike Stults
IRIS Data Management Center, Seattle Washington

Abstract
The IRIS Data Management Center (DMC) has served the seismological community for nearly 25 years. In that time we have offered data and information from our archive using a variety of mechanisms ranging from email-based to web applications and web services. Of these services, web services have quickly become the primary method for data extraction at the DMC. In 2011, the first full year of operation, web services accounted for 25% of the data shipped from the DMC. In 2015, over ~400 TB of data was delivered directly to users through web services, representing almost all of the data from the DMC that year. In addition to handling requests from a wide variety of researchers, the DMC switched all data extraction methods to use web services in 2014. On average the DMC now handles between 10 and 20 million requests per day delivered to web service interfaces.

The rapid adoption of web services is attributed to the many advantages they bring. For users, they provide on-demand data using an interface technology, HTTP, that is widely supported in nearly every computing environment and language. These characteristics, combined with readable and understandable data and extracted metadata facilitating integration of data access into existing workflows relatively easy. For the DMC, the web services provide an abstraction layer to internal repositories allowing for centralised optimisation of extraction workflow and easier evolution of these repositories. Lending further support to DMC’s push in this direction, the core services for station metadata, time series data and event parameters were adopted as standards by the International Federation of Digital Seismograph Networks (FDSN).

We expect to continue enhancing existing services and building new capabilities for this platform. For example, the DMC has created a federation system and tools allowing researchers to discover and collect seismic data from data centers federated using the FDSN standard web services. A capability in development, called Research-Ready Catalogue, will leverage the DMC’s quality assurance system to select data based on quality measurements. Within five years, the DMC’s web services have proven to be a robust and flexible platform that enables continued growth for the DMC. We expect continued enhancements and adoption of web services.

Key Advantages of Web Services

The DMC strives for efficiency in implementing a set of RISI-T web services that provide broad access to the data center’s holdings. These services perform a wide range of data extraction and processing techniques that include retrieving seismological time series and related metadata, formatting earthquake hypocentral parameters, evaluating instrument sensor response in common SEED RESP format and the returned data format.

- **Implementation:** Native to HTTP. This allows our web services to be adaptable to more complex network topologies.
- **Usability:** Provides easily readable, standardized documentation for end users. Everyone appreciates clear directions.
- **Development tools and clients:** Clients that retrieve and process data are sourced from the DMC, research community and widely available HTTP clients.
- **Data usage:** Provides data in useful formats. Formatting data easily consumed by human or machine means more time for processing and analysis.
- **Data access:** Standardized across global data centers.

Core Web Services from the DMC

The following web services represent the most persistently used services offered by the DMC, and the returned data format.

- **fdsnws-dateselect** Raw time series data in standardized miniSEED format
- **fdsnws-station** Station metadata in standardized StationXML and simple text format
- **irisws-timeseries** Event information in standard QuakeML and simple text format
- **irisws-event** Instrument sensor response in common SEED RESP format
- **irisws-resp** Instrument sensor response in common SAC, PAC, zeros format

A Variety of Data Request Methods

In addition to common HTTP clients such as wget, curl, and web browsers, a number of programs and toolkits have been developed to allow access to data from a variety of platforms.

command line, Fetch scripts
Java, IRIS-WS library
MATLAB, irisFetch
web, Wilber (event oriented)
Python, ObsPy
e-mail, BREQ, FAST

http://service.iris.edu/clients
http://service.iris.edu/clients
http://service.iris.edu/clients
http://ds.iris.edu/wilber/
http://docs.obspy.org/
http://ds.iris.edu/ds/nodes/dmc/manuals/breq_fast/

The IRIS Federator

The standardization and adoption of FDSN web services provides a common interface for seismological data access. Leveraging these common services, the DMC has built the IRIS Federator to help users discover and access data across data centers. The IRIS Federator allows simple, client-side federation and involves two steps: first the client sends the request to the IRIS-fedcatalog service, and then the client parses the catalog response and sends the request(s) to the identified data centers.

Data are requested by and returned directly to the client application from each federated data center. Data are not routed through a central location.

The main interface for the IRIS Federator is the irisws-fedcatalog web service:
http://service.iris.edu/irisws/fedcatalog/1/

Support for using the Federator to collect data from multiple centers is included in the DMC’s command-line FetchData and MATLAB irisFetch.m clients. An example:

```
$ FetchData -F +L -- -C LHZ --radius 3.5:96:70 -s 2004-12-26T00:03:00 -e 2004-12-26T02:00:00 -o fetch.mseed
```

http://ds.iris.edu/irisws/fedcatalog/1/

This work was supported through NSF awards: EAR-1261681, 1343709 & 1321260

[Image: Conceptual Flow chart illustrating how the IRIS Federator retrieves data from other FDSN data centers.

Client
Data Access Tool

1. Get channels

irisws-fedcatalog

2. Get data

Data center 1

Data center 2

Data center 3

The web services and data extraction methods developed by the IRIS Data Management Center.

[Diagram: Flowchart illustrating the process of data extraction using web services.]